DOCUMENT RESUME

ED 263 821 HE 018 865

AUTHOR King, Patricia M.; Kitchener, Karen Strohm
TITLE Reflective Judgment Theory and Research: Insights into the Process of Knowing in the College Years.
PUB DATE Mar 85
NOTE 46p.; Paper presented at the Annual Meeting of the American College Personnel Association (Boston, MA, March 24-27, 1985). Table 1 uses small print.
PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Abstract Reasoning; *Adult Development; *Cognitive Ability; College Students; Critical Thinking; *Educational Research; Higher Education; *Intellectual Development; Learning Theories; Persuasive Discourse; *Student Development
IDENTIFIERS *Reflective Judgment Model

ABSTRACT

The reflective judgment (RJ) model is described, along with research findings based on the model and contributions to understanding student intellectual development in higher education. The RJ model (Kitchener & King, 1981) describes a series of changes that occur in the ways adults/young adults understand the process of knowing, and subsequently, become better able to evaluate knowledge claims and defend one's own point of view on controversial views. Changes or stages of development in reasoning are described in terms of seven sets of assumptions about knowledge itself and how it is acquired. The longitudinal and cross-sectional research based on the model is reviewed, with attention to the effects of age, education, verbal ability, formal operational reasoning, critical thinking, scholastic aptitude, academic major, and sex on the development of RJ. The research results support the concept of RJ. While the data show a consistent pattern of upward shifts in RJ mean scores across age/educational groups and over time, the influence of age and education has not followed a consistent, predictable pattern. Although RJ is related to verbal ability and critical thinking, it is not simply a new description of either of these constructs. The use of a structured interview to assess reflective judgment level is described. An eight-page list of references and a three-page bibliography are included. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263821

Reflective Judgment Theory and Research:

Insights Into the Process of Knowing in the College Years

Patricia M. King
Department of College Student Personnel
Bowling Green State University

Karen Strohm Kitchener
School of Education
University of Denver

Presented at American College Personnel Meeting, Boston, 1985

Running Head: Reflective Judgment

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Patricia M King
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Abstract

The Reflective Judgment model (Kitchener and King, 1981) describes a series of changes that occur in the ways young adults and adults come to understand the process of knowing, and subsequently, to become better able to evaluate knowledge claims and defend one's own point of view on controversial views. This article reviews the cross-sectional and longitudinal research based on the model and examines the effects of age, education, verbal ability, formal operational reasoning, critical thinking, scholastic aptitude, academic major and sex on the development of Reflective Judgment, and shows how these findings contribute to our understanding of intellectual development beyond adolescence.

Reflective Judgment Theory and Research:
Insights Into the Process of Knowing in the College Years

Those who have the opportunity to work with college students can often identify changes that occur in students as they progress through college: students may seem more mature, more thoughtful, better organized, more sensitive and accepting toward those who are different from themselves, more confident in their own views, more aware of why others hold opinions with which they disagree, more articulate, etc. Often the changes are not so positive, however, and students may seem more defensive, less accepting of others, more rigid or dogmatic in their views, or more uncertain about how to evaluate the assertions or opinions of others.

While both positive and negative changes are often noticeable, the course, the direction and the causes of these changes are much more difficult to identify. Because student affairs is a profession committed to fostering the development of college students, questions of how and why students change are as important as whether they change. By being able to chart the course of developmental changes that occur in students and by being able to identify factors that either serve to promote or inhibit such development, student affairs practitioners can then make more informed decisions about ways of intentionally promoting student development.

Several models of intellectual development have attempted to chart the course of student development in this domain. When William Perry's (1970) model of the intellectual and ethical development of college students was first published, it was recognized as one of the

few available descriptions of the evolution of thinking after adolescence. Although others were also investigating young adult development at that time (Harvey, Hunt and Schroder, 1961; Kohlberg, 1969; Loevinger, 1966; Piaget, 1972), Perry's (1970) model was particularly stimulating for student affairs scholars and practitioners because it was explicitly grounded in observations of college students and how they understood and made meaning of their collegiate experiences. Several studies and programmatic applications have been based on this model, most notably Knefelkamp and Slepitza (1976), Mentkowski, Moeser and Strait (1983), Stonewater and Stonewater (1984) and Rodgers (1983). For a description of the model and a review of the early research, see King (1978); for a more recent review of Perry-based research, see Slepitza (1983).

In the last ten years there has been a growing interest in adult intellectual development and a dramatic increase in the number of theories or models that attempt to describe development in reasoning beyond adolescence (Arlin, 1975; Basseches, 1980, 1984; Broughton, 1975, 1978; Edelstein & Noam, 1982; Fischer, 1980; Gilligan & Murphy, 1979; Kitchener & King, 1981; Labouvie-Vief, 1982; Richards & Commons, 1984; Sinnott, 1984). This new generation of theoretical models has offered several alternative descriptions and explanations for observed changes in the ways young adults and adults reason about different kinds of problems in the cognitive domain. The Reflective Judgment model, the focus of this paper, is one such alternative description. While deeply influenced by Perry's work, this model has attempted to disentangle development of assumptions about knowledge from those of ethics and identity. Like Fischer (1980) and Broughton (1975), it

assumes these are different strands of development and that change across these areas does not necessarily occur in a lockstep fashion. In order to understand how these three aspects of the person are (or are not) related, we need to be able to describe and measure each separately. This article will describe the Reflective Judgment model, summarize the major findings of research conducted using the model, and how these findings contribute to our understanding of student intellectual development in higher education.

The Reflective Judgment Model

The Reflective Judgment model (Kitchener & King, 1981) describes a series of changes that occur in the ways young adults and adults understand the process of knowing. It is this process, it is argued, that allows a person to become better able to evaluate knowledge claims and to explain and defend one's own point of view on controversial issues. Specifically it describes the shifts that occur between adolescence and adulthood in people's assumptions about knowledge, about how one comes to know things, and how this affects the way a person justifies his or her beliefs or decisions.

These changes in reasoning are described in terms of seven distinct sets of assumptions about knowledge itself and how it is acquired. Each set of assumptions is characterized by its own logical coherency and constitutes a stage in the sense proposed by Glaserfeld and Kelley (1982): a period of time characterized by a "qualitative change that differentiates it from adjacent periods and constitutes one step on a progression" (p. 157). Each successive set of assumptions is posited to represent a more complex and effective form of justification and thus to provide a better means of evaluating and

defending a point of view. This practice of evaluating points of view pervades many adult roles and responsibilities. For example, deciding whether a particular sugar substitute is safe, voicing an opinion regarding what measures should be taken to balance the federal deficit, and selecting a new career all require a judgment to be made. These judgments may be better or poorer depending on the information and assumptions on which they are based. In problems such as these, the necessary information may be unavailable, it may be difficult to interpret, or the available reports may contain contradictory data or conclusions. Nevertheless people are called upon to make judgments and do so with varying degrees of comfort. Problems such as these have been termed "ill-structured" problems (Simon, 1978; Churchman, 1971; Wood, 1983) because they have no single correct answer. Thus the nature of the problem itself makes it more difficult to explain and defend one's strategy for arriving at an answer. As has been noted elsewhere (Kitchener, 1983; Kitchener & Kitchener, 1982; Wood, 1983; King, 1984) formal logic and deducation cannot alone account for reasoning well about ill-structured problems. Assumptions about how certain our knowledge is (e.g., absolute, probable) and what we can know (e.g., facts, interpretations) are critical in understanding the nature of ill-structured problems and forming a judgment about them. The Reflective Judgment model describes sets of assumptions about knowing that are used to make judgments about ill-structured problems, and how these assumptions evolve over time. The following stage descriptions are taken from the Reflective Judgment Scoring Rules (revised edition, 1984).

Stage 1 is characterized by a copy view of reality and knowledge, or the belief that there is an absolute correspondence between what is perceived and what is. Beliefs require little or no justification since it is assumed that one must only observe to know the truth. Knowledge and belief are indistinguishable; they simply exist. Since knowledge is absolute and predetermined, judgments which presume uncertainty are not made. Thinking appears naive and simplistic.

Stage 2 is characterized by the belief that there is a true reality that can be known with certainty, but which is not known by everyone. Certain knowledge is seen as the domain of authorities who are presumed to know the truth, e.g., scientist, teacher, or priest. When pre-existing answers do not already exist or cannot be learned by simple observation, beliefs are justified by copying the views of an authority. The admission that knowledge may not be directly and immediately known represents an advancement over stage 1. The individual moves from an egocentric, single-category belief system, "What I observe is true," to a two-category belief system: some beliefs are right and others are wrong.

The distinguishing feature of Stage 3 is the acknowledgement that truth is temporarily inaccessible because knowledge cannot always be immediately known, even to legitimate authorities. The belief that absolute truth will be manifest in concrete data at some future point is maintained. In the meantime, while evidence is incomplete, no one (including authorities) can claim to "know" beyond his or her own personal impressions or feelings. In other words, in areas of uncertainty, beliefs can be justified only on the basis of what "feels right" or what an individual wants to believe at that moment. Such

claims are basically a-rational. Stage 3 offers a further expansion of categories stage 2: right, wrong, and uncertain. Alternative conceptions of reality, diverse points of view or different theories are assimilated as areas of uncertain knowledge.

Stage 4 is characterized by the belief that reality cannot be known with certainty. The reasons given are concrete and specific, e.g., that an event happened too far in the past for us to have direct knowledge of it. Since there are many possible answers to every question, but no absolutely certain way to adjudicate between answers, knowledge claims are understood as being idiosyncratic to the individual. What is true for one person is then not necessarily true for the next. Stage 4 provides an advancement over Stage 3 in the admission that uncertainty is not just a temporary condition of the knowing process. The category "uncertain" has become further differentiated and generalized, and the former categories of right and wrong knowledge are seen as special exceptions rather than as general categories of knowing.

In Stage 5, the understanding that knowledge is uncertain generalizes to the belief that objective knowledge does not exist. Further, the belief is transferred to the realm of reality, with individuals claiming that reality can only be known through subjective interpretation of data. Beliefs are understood as strictly relative to a particular context or domain.

The logic is that beliefs are justifiable only within a given context because different contexts have different rules of inquiry and differing perspectives. For example, an individual might argue that given the perspective of a nuclear physicist, atomic energy plants

have an acceptable level of risk, but given the perspective of an ecologist, the risks are too high. A major advancement of stage 5 over stage 4 is the ability to evaluate an argument (e.g., the ecological view) by judging the relative merit of its component parts (i.e., the various pieces of evidence and their coherency). This allows the individual to move beyond the idiosyncratic justifications of stage 4 and to see justification as the process of making legitimate interpretations within a particular perspective.

Stage 6 is distinguished by the belief that since knowledge is based on subjective interpretations of data, objective knowledge is not possible to attain. Some judgments or beliefs may nevertheless be judged as better founded than others. This claim is based on principles of inquiry that generalize across contexts rather than on principles that are specific to a context. For example, adequacy of data and quality of arguments apply to both the nuclear physicists' position and the ecologists' position and allow for comparisions between them. Individuals begin to suggest that some points of view may be justified as better founded or more rational and reasonable than others. A major advancement of Stage 6 over Stage 5 is the ability to see relationships between subcategories in different domains. The individual has access to rules that allow him or her to make evaluations across these subcategories, e.g., choosing the position that has the more coherent argument or the less drastic consequences.

Stage 7 is characterized by the belief that while reality is never a "given", facts and interpretations of those facts can be constructed into evaluated knowledge claims about reality. Knowledge must be constructed via critical inquiry or through the synthesis of

existing views and evidence into more cohesive and coherent explanations. While the knower is limited in perceiving reality directly, it is possible through the rules of critical inquiry or synthesis to determine that some judgments or interpretations have greater truth value than others and/or to suggest that a judgment can be justified as the most reasonable current solution to a problem. Subjects exhibit a complex understanding of issues and an awareness that in the future, with the collection of additional data or its reinterpretation, they may need to reformulate their points of view.

Table 1 summarizes three dimensions of the Reflective Judgment model, how certain knowledge is, how knowledge is gained and how beliefs are justified. These characteristics illustrate how the

Insert Table 1 about here

different sets of assumptions about knowledge and its acquisition are evidenced when one observes the reasoning of young adults and adults.

The Reflective Judgment Interview

Reflective Judgment level is assessed through a structured interview called the Reflective Judgment Interview (RJI). Subjects are asked to respond to a specific set of questions about four ill-structured problems from the physical and social sciences and covering the areas of historical knowledge, scientific knowledge, knowledge of current events and the relationship between religion and science. The interview takes 45-60 minutes to complete.

The interview was designed to meet several criteria, first, that the content of the interview be within the intellectual domain and

sample individuals' assumptions about knowledge across several content areas. The intent was to differentiate intellectual issues from identity or ethical issues. We assumed that if we could clearly measure and understand how people reasoned and drew conclusions in the intellectual domain, we could later more clearly reexamine the relationship between intellectual decision making and ethical and/or identity development.

The second criterion was that the measurement format not be limited to use in formal educational settings or with college students. To tie the content or scoring of the interview to issues or problems that are unique to the college campus would disallow asking certain sets of questions, such as the following: Do high school students use different sets of assumptions than do college students? Do individuals change their assumptions between high school and college? Do young adults who have not participated in a college experience show the same sequence of development as do college students? Does development continues to take place in the middle adult years when people are no longer in college? In this way, the RJI was designed to allow the model to be tested over a broad range of subjects between the adolescent and middle adult years. To date, it has been successfully used with subjects ranging in age between 16 (Kitchener & King, 1981) and 55 (Glatfelter, 1982), with undergraduate and graduate college and university students (Brabeck, 1983; Griffith, 1984; Hayes, 1981; Kitchener & King, 1981; Mines, 1980/1981; Shoff, 1979; Schmidt, 1983; Strange & King, 1981; Welfel, 1980; Wood, 1980), with non-student adults (Glatfelter, 1982; King, Kitchener, Davison, Parker & Wood, 1983; Kitchener, King & Wood, 1984; Lawson, 1980; Schmidt, 1983) and with German university students (Kitchener & Wood, 1983).

The third design criterion was that the interview format and rating procedure be standardized. Hence, a rater and interviewer certification program has been established. Virtually all of the studies reviewed here used RJI's conducted by certified reflective judgment interviewers and rated by certified raters. The suggested procedure is for the interviews to be tape recorded and transcribed, and then for each dilemma to be rated independently by two certified raters. Mean scores are derived by averaging ratings by both raters across all dilemmas. These are used for group comparison purposes. The patterns in scores for individuals have been examined using Davison's (1979) test for sequentiality, a method for determining whether patterns of stage usage across adjacent and non-adjacent stages occurs in a predictable fashion. These analyses are reported by King et al. (1983) and Kitchener and Wood (1983). Kitchener and King (1981) also examined individual change scores by noting which stages decreased and which stages increased in usage over time.

The psychometric properties of the RJI have been recently reviewed in detail elsewhere (Mines, 1982). In general, inter-rater reliability has been moderate to high, depending on the heterogeneity of the sample tested, and inter-rater agreement for first round ratings (the most conservative index) has consistently ranged between 70% and 80%. Test-retest reliability on a homogeneous sample over a three month period has ranged from .71 to .83 (Sakalys, 1982). Cronbach's alpha, a measure of internal consistency, estimates the extent to which the dilemmas are measuring the same construct. This has ranged from .62 (Welfel, 1982) to .96 (Kitchener & King, 1981). Again, the heterogeneity of the samples used influences the size of Cronbach's alpha.

It should be noted that the RJI is an experimental procedure that was designed to yield information about people's reasoning processes in a rich and complex way. While it has been a useful tool for theory development, it is nevertheless time consuming and expensive to use. A paper and pencil form of the RJI is currently under construction (King, 1983).

Summary of Research Findings

This section will review and summarize the major findings that have occurred to date in studies using the Reflective Judgment model. These findings will be organized around four major research questions.

Does Development in Reflective Judgment Occur in the Predicted Fashion? One of the first questions that must be addressed in establishing the validity of any hierarchical, developmental model is whether people change over time in the manner predicted by the model. One of the first predictions tested using the Reflective Judgment (RJ) model was that the assumptions of older, more highly educated subjects would be most similar to the more advanced stages in the model, and that those of younger, less educated subjects would reflect the assumptions of earlier stages. In other words, older, more highly educated subjects would be expected to earn higher RJ scores than younger subjects with less education. It was assumed that age alone would not necessarily lead to higher scores because there is no reason to believe that the passage of time by itself should influence developmental change unless it corresponds with some other intervening event (e.g., maturation of neurons, involvement in stimulating educational environments, etc.). Since educational environments focus on the acquisition and utilization of knowledge, changes in assumptions about knowledge should most

likely occur with people are involved in such environments. The failure to find differences over age/educational levels would seriously challenge the validity of the model. Table 2 shows the mean RJ scores by educational level for twelve studies using traditionally aged students. Mean scores correspond to stages, i.e., a mean score of 2.0 shows that the average score for a given sample was Stage 2. It should be noted that in some studies students were selected by age and educational level and in others they were randomly selected within an educational level. In the latter case, while the modal age of the sample was age typical, the samples may have included older adults who were enrolled in traditional college programs.

Insert Table 2 about here

As was predicted, RJ scores of these traditionally aged students increased significantly with educational level in virtually every study reported. While specific means differed for different samples and institutions, small positive differences of about one-half stage were found between the freshman and senior years of college and of one to two stages between the junior or senior year of college and advanced graduate study. The mean scores for each educational level across studies reflect a distinct upward trend in RJ scores over age and educational levels. However, because all of these studies used cross-sectional designs, the differences observed between groups could be attributed to cohort differences or to differences in the historical influences each group experienced (e.g., the Vietnam War, the Watergate scandal, etc.). Longitudinal data are important for a more accurate understanding of individual change over time.

Five longitudinal studies have been completed to date (Brabeck & Wood, 1983; King et al., 1983; Kitchener et al., 1984; Schmidt, 1983; Welfel & Davison, 1983) and have supported the trends found in the cross-sectional studies (see Table 3). With one exception (Schmidt's non-traditional aged college freshmen), group mean scores increased significantly for all groups tested at one to four year intervals. These consistent upward shifts in mean scores over time, combined with

Insert Table 3 about here

consistent group differences found in the cross-sectional studies, offer an affirmative answer to the question of whether groups of people change over time in the manner predicted by the model. (See Brabeck, 1984, for a review of longitudinal studies of adult intellectual development, and a discussion of the limitations of longitudinal designs.)

The next question to be addressed is whether all individuals tested in the longitudinal studies changed in the predicted fashion. King et al. (1983) report that while the mean scores of all of their former high school subjects showed upward change over a two-year period, this was true of only 64% of the former college juniors and 40% of the former graduate students. Eleven percent of the former college students and 20% of the former graduate students showed downward changes in mean scores. The findings of Brabeck and Wood (1983) and Schmidt (1983), for one to three year periods, showed a similar pattern in change scores. By contrast, for longer time periods (4 to 6 years), Welfel and Davison (1983) and Kitchener et al. (1984) report that 92-100% of their college samples showed upward change.

Perhaps individuals whose scores did not increase over the shorter time period were proceeding at a slower pace and may have been preoccupied with other developmental tasks at that point in their lives.

Age and educational level have been confounded in the studies reviewed thus far, and several researchers have attempted to discover whether educational level or age (where age represents a correlate of other life experiences) has a more potent influence on Reflective Judgment development. The question was first investigated through quasi-experimental studies with samples that differed in age and/or educational level and more recently in a longitudinal study.

What are the Relative Effects of Age and Education on the Development of Reflective Judgment? The first studies addressing this question (Shoff, 1979; Strange & King, 1981) supported the assumption that development in Reflective Judgment is more closely related to education than to other life experiences. Shoff (1979) compared four groups of students, two groups of seniors (traditional and non-traditional ages) and two groups of students in a continuing education program who had no previous college experience (traditional and non-traditional ages). She found that the seniors (taken as a group) scored significantly higher than the entry level continuing education students, and that the older continuing education students scored higher (but not significantly so) than their younger counterparts.

Strange and King (1981) tested four groups of college students matched on scholastic aptitude test score. The groups included 18 and 22-year-old freshmen and 22 and 26-year-old seniors. They found

significant differences in RJ scores by class (i.e., between seniors and freshmen) but not by age. They concluded that formal educational experiences appear to play a more important role in the development of Reflective Judgment than do general life experiences.

Subsequent cross-sectional studies were less clear about the role of education in development. Lawson (1980) compared two groups of graduate students (beginning and advanced) and two groups of non-student college graduates who were matched with the graduate student sample on age and scholastic aptitude. This design allowed her to compare the relative influence of three variables on intellectual development: selection (comparing those who choose to further their formal education and those who do not); maturation (comparing subjects in different age groups); and education (comparing those with different educational levels). She found that graduate students scored significantly higher than non-students, reflecting the influence of the selection factor. The education factor was not found to be significantly related to RJ development, and the age factor did not remain significant when the effects of verbal aptitude and socio-economic status were partialled out. However, the degree of difference between the scores of the older and younger graduate students was not significantly greater than between the older and younger non-students. She points out that the failure to find a significant effect due to education may not be a consequence of an "absence of any educational effect, but to an inconsistency in effect" (p. 68). For example, the scores of her advanced graduate students were bi-modal, with 30% of the sample earning a mean score of 4.0 and 40% earning a mean score of 5.5. An additional 15% scored at 6.0. Thus, it appears that education promoted

development for one subset of this graduate sample, but not for another. She also noted that the means for her non-student groups (4.0 and 4.25) were similar to those reported earlier for college seniors. It appeared that their scores had not increased since college, suggesting that intellectual development may tend to level off after college graduation.

The results of a study by Schmidt (1983) are also provocative. In a design similar to that used by Strange and King (1981), she compared traditional aged freshmen (mean age, 19.7 years) with non-traditional aged freshmen (mean age, 22.9) and traditional aged juniors with a mean age of 21. She found that age and educational factors in combination are related to the development of Reflective Judgment. For example, the junior class scored significantly higher than the freshmen class (when the traditional and non-traditional groups were combined). This design, however, confounds age and education. Controlling for the age factor by comparing the traditional juniors and non-traditional freshmen (both age 21), no effect was found for educational level. Controlling for the education factor by comparing the traditional and non-traditional freshmen, no effect was found for age. Breaking this analysis down by sex, however, shows that the older women scored significantly higher than younger women, regardless of years in school. This suggests that age (or more precisely, those experiences that are associated with age) appears to be a more potent variable for women's intellectual development than education. The same pattern was not found for men.

In a longitudinal follow-up of the two freshmen groups four years later, Schmidt (1983) found a significant increase in scores. However,

the traditionally aged students accounted for the differences, since there was virtually no increase in mean scores for the non-traditional group.

Glatfelter (1982) also found a stronger effect for age than education in a study of traditional and non-traditional women. She compared traditionally aged first year college women and fourth year college women with re-entry women in their first and fourth year of college. She also compared a sample of older women who met the criteria for acceptance to college but who had not re-entered; these subjects were matched on age to the re-entry women. She found a significant effect for age, but not for educational level for the total sample; she also found a significant difference between the freshmen and seniors for the traditionally aged group.

Averaging the Reflective Judgment scores across studies, the mean score for non-traditionally aged first year college students (n = 94) is 3.57, which is remarkably similar to the mean of the traditionally-aged freshmen ($\bar{x}$ = 3.60, n = 216) reported in Table 2. In other words, on the average, re-entry students look remarkably like traditional college freshmen in terms of their assumptions about the process of knowing, and not like college seniors or graduate students.

What can we conclude from this group of studies about the role that general life experiences (age) and education play in the development of Reflective Judgment? Clearly, design and sampling problems affect one's interpretation of this group of studies. The issue of cohort differences is particularly critical for the cross-sectional studies, as younger subjects may be different than older peers when they reach their age. This is particularly true in evaluating Glatfelter's

(1982) findings, since the women in her study were 8 to 37 years older than her 18-year-old freshmen, and there is no evidence of what their RJ scores were when they left high school. Conceivably, they might have reached their current level of RJ before they left high school. While this may seem improbable, it should be noted that Kitchener and Wood (1984) found that beginning German college students (just out of secondary school) scored as high on RJ as have seniors in samples of American college students and Kitchener et al. (1984) reported significant cohort differences in their samples of American students.

Because of possible cohort differences confounding results, longitudinal studies of subjects at the same age who do and do not go on to college are particularly helpful in sorting out the effects of age and education. In the one known study of such groups, Kitchener et al. (1984) report that while subjects who do not go on to college show some increase in RJ scores, those who earned college degrees have significantly larger gain scores.

Based on the above studies and the longitudinal studies reported in the prior section, some tentative conclusions about the role of general life experiences and education may be drawn. First, it appears that age and education taken together make a difference in promoting intellectual development as described by the Reflective Judgment model. The pace of change, however, varies greatly across individuals, with some experiencing dramatic change over short periods, and some experiencing little if any change. On the average, RJ level tends to increase about one-half a stage over the four years of college for traditionally aged college students. While this degree of change may not seem dramatic, the fact that it also reflects a shift to a style

of reasoning that explicitly uses evidence in making judgments is dramatic. For without this fundamental component, a rational approach to problem solving is impossible. Traditionally aged college students tend to score in the middle RJ stages, even as seniors. Reasoning at the higher RJ stages is rarely observed among undergraduate students.

Second, for traditionally aged undergraduates, education appears to play a stronger role in the development of Reflective Judgment than do other life experiences outside of the university. Other life experiences may also play a role in promoting development and this appears to be particularly true for re-entry women. By contrast to studies of traditionally aged students, both Glatfelter's (1982) and Schmidt's (1983) findings would support the contention that undergraduate education is not effective in promoting development of these women to higher levels of Reflective Judgment. These findings support a suggestion made by Kitchener (1983) that undergraduate education is particularly suited to promote development on the lower and middle Reflective Judgment stages, since it encourages the examination of multiple perspectives, but not to higher levels, where well-defended points of view are required. Further, while other life experiences may promote development for some women, this was not true of the men in Schmidt's sample nor of the older men and women in Shoff's (1979) or Strange and King's (1981) samples. It may be that for some women the decision to re-enter school involves considering and examining alternative conceptions of women's roles; this process itself may promote the development that typically takes place during the undergraduate years.

Third, taken singly, neither life experience nor undergraduate education appears to promote development to the highest level of

Reflective Judgment. While graduate education may be critical in promoting development to higher stages, its effect is inconsistent. Last, the undergraduate experience may effect traditionally aged and non-traditionally aged men and women students differently. It appears that researchers studying Reflective Judgment should not assume that these groups are the same and should differentiate between them in sampling.

What is the Relationship Between Reflective Judgment and Other Related Constructs? It is important to ask of any new model of development whether it is a new description of a previously researched construct. This is particularly important for a model that purports to measure an aspect of intellectual development, since the intelligence testing, formal operations, and critical thinking traditions already have established research literatures.

Prior longitudinal research on intellectual development (Bayley & Oden, 1955; Honzik and MacFarlane, 1973; Kansas and Bradway, 1971) support the observation that verbal ability (a measure of verbal intelligence that is strongly related to learning) increases with age at least until the mid-adult years. It is particularly important to take this finding into account when investigating a developmental construct that is measured using an interview format. Several Reflective Judgment studies have used Terman's (1973) Concept Mastery Test (CMT) to estimate the relationship between verbal ability and Reflective Judgment score. As with correlations between any two variables, the size of the correlation is affected by the heterogeneity of the sample.

Correlations between Reflective Judgment score and the CMT have consistently been positive, and with the exception of the very hetero-

geneous Kitchener and King study (1981), they have been low to moderate, suggesting that verbal ability and Reflective Judgment are different constructs. In four studies (Kitchener & King, 1981; Glatfelter, 1983; Lawson, 1980; Schmidt, 1983) of cross-sectional samples, however, differences between the RJ scores of one or more groups could be statistically accounted for by differences in verbal ability.

The longitudinal relationship between the two measures provides a stricter test of the hypothesis that verbal ability can account for differences in Reflective Judgment scores. With longitudinal data, subjects act as their own control, thus the failure to find differences in RJ score between different groups at two points of time when controlling for verbal ability cannot be attributed to cohort differences. In fact, in the three longitudinal studies that have measured both Reflective Judgment and verbal ability (King et al., 1983; Schmidt, 1983; Welfel & Davison, 1983), differences in a subject's Reflective Judgment score over time could not be statistically accounted for by changes in verbal ability. These data support the claim that Reflective Judgment is related to but not the same construct as verbal ability.

A second related construct that has been investigated is formal operational reasoning (Inhelder & Piaget, 1958). Since formal operations require a high level ability to see logical relationships, it might be that Reflective Judgment is measuring these logical relationships in new domains (e.g., social science or history). King (1977) investigated the relationship between Reflective Judgment and two formal operations tasks, the chemicals and the pendulum tasks. Despite the fact that 91% of the sample scored as fully formal on either one or both tasks, significant differences were found between high school,

college, and graduate students on RJ scores. The correlations between the two Piagetian tasks and Reflective Judgment scores were low and nonsignificant. While the low correlations can be accounted for, in part, by the limited range of scores on the Piagetian tasks, King concluded that the observed differences between groups in RJ scores could not be attributed to differences in formal operations. Formal reasoning may be a necessary but not sufficient skill for the attainment of advanced levels of Reflective Judgment. Kitchener and Kitchener (1981) have offered a conceptual explanation of the differences between the two models, pointing out that logic alone cannot account for differences in assumptions about knowledge and the ability to construct probabilistic arguments. In a recent review of formal reasoning in adults, King (1984) has also noted that many who criticize the limits of formal logic in adequately describing adult thought do so because of the nature of the questions posed in formal reasoning tasks, questions that require a simpler level of cognitive processing (Kitchener, 1983) than those that fully reflect the complexity of mature adult thought.

Brabeck (1983), Mines (1980/1981) and Wood (1980) have noted the similarities in the definitions of critical thinking and Reflective Judgment. Brabeck (1983) found correlations ranging from .37 to .48 between the Watson Glaser Critical Thinking Appraisal (WGCTA) and RJ. In addition, she found that while individuals who scored low on critical thinking had consistently low RJ scores, high critical thinkers had more varied RJ scores. Wood (1980) evaluated the necessary but not sufficient relationship between Reflective Judgment score and the subtests of the WGCTA and the Cornell Critical Thinking Test (CCTT). He found that two subtests (WGCTA - Recognition of Assumptions and

CCTT - Choosing Useful Hypothesis - Testing Predictions) performed in a way which one would expect if high levels of critical thinking were necessary but not sufficient for high levels of Reflective Judgment. Mines (1980/1981), using a sample of college freshmen, seniors and graduate students, found that the critical thinking skills of interpretation (WGCTA), detection of fallaciously ambiguous arguments (CCTT), deduction (WGCTA), and inference (WGCTA) could distinguish between those who scored at Reflective Judgment stage 3 or below and those who scored above stage 3. Based on these studies, it appears that the WGCTA taken as a whole or the subtests of the WGCTA and the CCTT cannot account for the stage differences in Reflective Judgment, particularly above stage 3, where the role of evidence in constructing arguments becomes explicit.

Investigators have also attempted to rule out several potentially confounding variables as an explanation for group differences found in Reflective Judgment scores. Kitchener and King (1981) found a significant correlation (r = .40, p < .001) between RJ scores and verbal fluency (i.e., number of words a subject spoke during the RJ interview), and between RJ score and socio-economic status (r = .22, p < .05). They reported, however, that controlling for the effects of verbal fluency and socio-economic status added little to the effects of controlling for verbal ability (CMT) alone and could not account for group differences in RJ. Lawson (1980) and Welfel (1979) found a similar pattern of results when controlling for socio-economic status and CMT score. Other investigators have controlled for scholastic aptitudes usually using either SAT or ACT scores. Correlations have been low to moderate, ranging from .15 to .56 (Brabeck, 1983; Hayes, 1981; Kitchener, 1977/1978;

Sakalys, 1982; Welfel, 1982) and scholastic aptitude test scores have not statistically accounted for group differences in Reflective Judgment scores (Mines, 1982; Welfel, 1982).

In summary, the evidence suggests that Reflective Judgment is not the same as verbal ability, critical thinking, formal operations, verbal fluency, socio-economic status or scholastic aptitude. In other words, the aspects of thinking that the Reflective Judgment model is measuring cannot be attributed to any one of these constructs. It should be noted, however, that this does not mean that Reflective Judgment describes "intellectual" development in and of itself. It is more accurate to suggest that it describes one aspect of intellectual development, i.e., development of assumptions of knowledge and their impact on judgment, and that this development may be affected by development in other aspects of the intellect.

Are Academic Major and Sex Related to the Development of Reflective Judgment? Researchers have attempted to answer one other major question about the development of Reflective Judgment: Do different groups of subjects (e.g., males-females, liberal arts or engineering majors) score differently on RJ? The question underlying this group of studies is whether the life experiences associated with particular groups appear to promote or retard development. This review includes only those studies that used the standard Reflective Judgment dilemmas and that reported the educational level of the subjects.

Six studies (Brabeck, 1980; King & Parker, 1978; King, Wood & Mines, 1984; Schmidt, 1983; Welfel, 1982; Welfel & Davison, 1983) have investigated the effects of college major on RJ score. Brabeck (1980) found a non-significant correlation between RJ and number of courses

taken in different disciplines. In the other studies, undergraduate students in humanities or liberal arts majors were compared to majors in more technically oriented fields, e.g., engineering or agriculture. No significant differences in RJ were found in any study that could be attributed to major alone. This was true both in the cross-sectional and longitudinal studies. By contrast, King et al. (1984) found that graduate students in social science majors scored significantly higher than did students in mathematical sciences. This finding may reflect a selection effect or it may suggest that the differential effects of major on ways of knowing are only apparent after more intense periods of study. Some specific college experiences have also been associated with higher Reflective Judgment scores (Schmidt, 1983; Welfel & Davison, 1983). For example, Welfel and Davison (1983) suggest that students who live "multi-dimensional" lives (i.e., those who are involved in athletics, religious and extracurricular activities, etc.) have higher Reflective Judgment scores.

The effects of gender on RJ score have been evaluated in most studies. The results of these studies taken individually have been conflicting. King and Parker (1978), Kitchener and King (1981), Welfel (1982), and Welfel and Davison (1983) found no significant differences in Reflective Judgment score by sex. Significant sex effects have been found in other studies and in the majority of cases (Lawson, 1980; King, et al. 1984; Shoff, 1979; Strange & King, 1981) males scored higher than females. Schmidt (1983), however, found at her first testing that the traditionally-aged female juniors and nontraditionally-aged female freshmen scored higher than their male counterparts. The sex differences in these studies were not statistically

attributable to verbal ability or scholastic aptitude, although the failure to find sex differences in some studies may be attributable to verbal ability. For example, in the two-year and six-year follow-up of the Kitchener and King (1981) and King and Parker (1978) studies, King et al. (1983) and Kitchener et al. (1984) found a group by sex interaction after covarying CMT scores. Despite having higher CMT scores than men and despite the positive correlation between CMT and RJ scores, women's RJ scores were not consistently higher than men's across all groups.

Table 4 reports the mean RJ scores by sex and educational level for those studies that do not involve retesting of the same subjects

Insert Table 4 about here

longitudinally. The overall pattern of sex differences reveals that males tend to score higher than females. Although the differences in most cases are small, they tend to increase with age and education, possibly suggesting that higher education does not have as strong of an effect on women's judgment and assumptions about knowing as it does on men's and that the effect may be cumulative. Welfel (1982) pointed out that significant sex differences have been found predominently in studies that include samples of nontraditional students. Although the sex differences cannot be attributed just to differences in these groups, it underlies the need to differentiate traditional and non-traditional students by sex in future RJ research.

Discussion

As Cross (1981) has so astutely noted: "The systematic accumulation of knowledge is essential for progress in any profession.

In an applied profession, however, theory and practice must be constantly interactive. Theory without practice is empty, and practice without theory is blind" (p. 110). The Reflective Judgement model described here evolved out of the collective concerns of many people who were struggling to understand the changes they observed in college students. It also evolved out of a collective curiousity about how adults come to grips with the controversial issues of the day. Informal insights and prior information on these issues had become insufficient; the systematic accumulation of new knowledge was required. The studies reviewed here have added to our knowledge base about one aspect of intellectual development. What have we learned?

First, we have learned that adults do change in the assumptions they hold about knowledge itself and how it is acquired, and that these assumptions are reflected in the way they resolve controversial issues. Development towards more adequate ways of resolving such issues appears to evolve in a manner consistent with the posited sequence of Reflective Judgment stages. The cross-sectional and longitudinal data show a consistent pattern of upward shifts in RJ mean scores across age/educational groups and over time.

Second, both age and education have emerged as important factors related to the development of Reflective Judgment, but their influence does not follow a consistent, predictable pattern. A finer-grained analysis of specific types of experiences within educational and general life experiences is needed in order to more fully understand this issue. This is particularly important as more institutions attempt to attract and meet the needs of non-traditional students since the effectiveness of the collegiate experience in promoting reflective thinking seems to be so variable among these students.

Third, the Reflective Judgment construct has emerged as a description of an aspect of intellectual development that is distinct from other descriptions. While it is clearly related to both verbal ability and critical thinking, the studies reviewed here show that it cannot be considered to be simply a new description of either of these constructs.

The fourth major question in this review looked at the effects of academic major and sex on Reflective Judgment. Academic major has not emerged as a powerful factor in understanding RJ development in the undergraduate years. Sex differences across studies, however, have been provocative, if for no other reason than the inconsistency of their appearance. Again, this factor appears to be a more potent one for non-traditional students.

These studies have contributed to our general knowledge base for understanding changes that occur in the way people reason about controversial problems. The major issues of our day require clear thinking and taking a stand, even when complete and unambiguous information is simply not available. Assisting college students, our colleagues, other decision-makers and even ourselves to do so in a more effective way is a natural goal to work toward as an outcome of this line of research. As Parker (1977) has noted, ". . . the nature of theory is such that it does not lead directly to practice, and the nature of practice is such that it does not proceed without theory" (p. 420). Research based on the Reflective Judgment model by itself does not lead directly to practice. But it does provide a foundation for understanding one aspect of development in the young adult and adult years, a foundation on which subsequent translations into practice can be based.

REFERENCES

Arlin, P.K. (1975). Cognitive development in adulthood: A fifth stage? Developmental Psychology, 11, 602-606.

Basseches, M. (1980). Dialectical schemata: A framework for the empirical study of the development of dialectical thinking. Human Development, 23, 400-421.

Basseches, M.A. (1984). Dialectical thinking as a metasystematic form of cognitive organization. In M.L. Commons, F.A. Richards and C. Armon (Eds.), Beyond formal operations: Late adolescent and adult cognitive development (pp. 216-238). New York: Praeger.

Bayley, N. & Oden, M.H. (1955). The maintenance of intellectual ability in gifted adults. Journal of Gerontology, 10, 91-107.

Brabeck, M. (1980). The relationship between critical thinking skills and the development of reflective judgment. (Doctoral dissertation, University of Minnesota, 1980.) Dissertation Abstracts International, 41, 4647A.

Brabeck, M. (1983). Critical thinking skills and reflective judgment development: Redefining the aims of higher education. Journal of Applied Developmental Psychology, 4, 23-34.

Brabeck, M. (1984) Longitudinal studies of intellectual development during adulthood: Theoretical and research models. Journal of Research and Development in Education, 17, 12-27.

Brabeck, M. & Wood, P.K. (1983). A longitudinal study of well and ill-structured problem solving in college age women. Unpublished manuscript.

Broughton, J.M. (1975). The development of natural epistimology in adolescence and early adulthood. Unpublished doctoral dissertation, Harvard University.

Broughton, J.M. (1978). Development of concepts of self, mind, reality, and knowledge. New Directions for Child Development, 1, 75-100.

Churchman, C.W. (1971). The design of inquiring systems: Basic concepts of systems and organizations. New York: Basic Books.

Cross, K.D. (1981). Adults as learners. San Francisco: Jossey-Bass.

Edelstein, W. & Noam, G. (1982). Regulatory structures of self and 'postformal' stages in adulthood. Human Development, 25, 407-422.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, 87, 477-531.

Gilligan, C. & Murphy, J.M. (1979). Development from adolescence to adulthood: The philosopher and the dilemma of the fact. New Directions for Child Development, 5, 85-99.

Glaserfeld, E. von & Kelley, M.F. (1982). On the concepts of period, phase, stage, and level. Human Development, 25, 152-160.

Glatfelter, M. (1982). Identity development, intellectual development, and their relationship in reentry women students. (Doctoral dissertation, University of Minnesota, 1982. Dissertation Abstracts International, 43, 3543A.

Griffith, J. (1984). [Reflective Judgment scores, Davidson college students]. Unpublished raw data.

Harvey, O.J., Hunt, D.E. and Schroder, H.M. (1961). Conceptual systems and personality organization. New York: John Wiley.

Hayes, A.B. (1981). An investigation of the effects of dilemma content on level of reasoning in the reflective judgment interview. (Doctoral dissertation, University of Utah, 1981. Dissertation Abstracts International, 40, 2564B.

Honzik, M.P. & MacFarlane, J.W. (1973). Personality development and intellectual functioning from 21 months to 40 years. In L.F. Jarvik, C. Eisdorfer, & J.E. Blum (Eds.), Intellectual functioning in adults. New York: Springer.

Inhelder, B. & Piaget, J. (1958). The growth of logical thinking. London: Routledge & Kegan Paul.

Kansas J. & Bradway, K. (1971). Intelligence at middle age: A 38-year follow-up. Developmental Psychology, 5, 333-337.

King, P.M. (1977). The development of reflective judgment and formal operational thinking in adolescents and young adults. (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 38, 7233A.

King, P.M. (1978). William Perry's theory of intellectual and ethical development. In L.L. Knefelkamp, C. Widick & C.A. Parker (Eds.), New directions for student services, 4, 35-51.

King, P.M. (1983). Reflective judgment questionnaire, Technical Report #1. Unpublished manuscript.

King, P.M. (1984). Formal reasoning in adults: A review and critique. In R.A. Mines and K.S. Kitchener (Eds.), Social-cognitive development in young adults. Praeger Press.

King, P.M., Kitchener, K.S., Davison, M.L., Parker, C.A. & Wood, P.K. (1983). A longitudinal study of reflective judgment and verbal aptitude in young adults. Human Development, 26, 106-116.

King, P.M. and Parker, C.A. (1978). Assessing intellectual development in the college years. Unpublished manuscript, Instructional Improvement Project, 1976-1977, University of Minnesota, Minneapolis.

King, P.M., Wood, P.K. & Mines, R.A. 1984). Critical thinking by college and graduate students: Differences across measures and majors. Unpublished manuscript.

Kitchener, K.S. (1977/1978). Intellectual development in late adolescents and young adults: Reflective judgment and verbal reasoning. (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 39, 936B.

Kitchener, K.S. (1983). Educational goals and contemporary models of reflective thinking. Educational Forum, Fall 1983, 75-95.

Kitchener, K.S. & King, P.M. (1981). Concepts of justification and their relationship to age and education. Journal of Applied Developmental Psychology, 2, 89-116.

Kitchener, K.D., King, P.M. & Wood, P.K. (1984, August). A longitudinal study of epistemic cognition in young adults. Paper presented at the meeting of the American Psychological Association, Toronto, Ontario.

Kitchener, K.S. & Kitchener, R.F. (1982). The development of natural rationality: Can formal operations account for it? In J. Meacham & N.R. Santilli (Eds.), Social development in youth: Structure and content. Basel: Karger.

Kitchener, K.S. & Wood, P.K. (1984). Development of concepts of justification in German university students. Unpublished manuscript.

Knefelkamp, L.L. & Slepitza, R.A. (1976). A cognitive-developmental model of career developemnt - An adaptation of the Perry scheme. The Counseling Psychologist, 6, 53-58.

Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D.A. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally.

Labouvie-Vief, G. (1982). Dynamic development and mature autonomy. Human Development, 25, 161-191.

Lawson, J.M. (1980). The relationship between graduate education and the development of reflective judgment: A function of age or educational experience. (Doctoral dissertation, University of Minnesota, 1980). Dissertation Abstracts International, 41, 4655A.

Loevinger, J. (1966). The meaning and measurement of ego development. American Psychologist, 21, 195-206.

Loevinger, J. (1976). Ego development: Conceptions and theories. San Francisco: Jossey Bass.

Mentkowski, M., Moeser, M. & Strait, M. (1983). Using the Perry scheme of intellectual and ethical development as a college outcomes measures: A process and criteria for judging student performance. Milwaukee, Wisconsin: Alverno College Productions.

Mines, R.A. (1982). Student development assessment techniques. In U. Delworth & G.R. Hanson (Eds.), New directions for student services, 20, 65-92.

Mines, R.A. (1980/1981). An investigation of the developmental levels of reflective judgment and associated critical thinking skills. (Doctoral dissertation, University of Iowa, 1980). Dissertation Abstracts International, 41, 1495A.

Parker, C.A. (1977). On modeling reality. Journal of College Student Personnel, 18, 419-425.

Perry, W. Jr. (1970). Forms of intellectual and ethical development in the college years. New York: Holt, Rinehart and Winston.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. Human Development, 15, 1-12.

Richards, F.A. and Commons, M.L. (1984). Systematic, matasystematic, and cross-paradigmatic reasoning: A case for stages of reasoning beyond formal operations. In M.L. Commons, R.A. Richards, and C. Armon, Beyond formal operations: Late adolescent and adult cognitive development, New York: Praeger.

Rodgers, R.F. (1983). Using theory in practice. In T.K. Miller, R.B. Winston, Jr. & W.R. Mendenhall (Eds.), Administration and leadership in student affairs: Actualizing student development in higher education. Muncie, Indiana: Accelerated Development, Inc.

Sakalys, J.A. (1982). Effects of a research methods course on nursing students' research attitudes and cognitive development. (Doctoral dissertation, University of Denver, 1982). Dissertation Abstracts International, 43, 2254A.

Schmidt, J.A. (1983). The intellectual development of traditionally and nontraditionally aged college students: A cross sectional study with longitudinal follow-up. (Doctoral dissertation, University of Minnesota, 1983). Dissertation Abstracts International, 44, 2681A.

Shoff, S.P. (1979). The significance of age, sex, and type of education on the development of reasoning in adults. (Doctoral dissertation, University of Utah, 1979). Dissertation Abstracts International, 40, 3910A.

Sinnott, J.D. (1984). Postformal reasoning: The realitivistic stage. In M.L. Commons, F.A. Richards, & C. Armon (Eds.), Beyond formal operations: Late adolescent and adult cognitive development. New York: Praeger.

Slepitza, R.A. (1983). Commitment within the Perry scheme: A question of structural change. (Doctoral dissertation, University of Maryland, 1983).

Stonewater, J.K. & Stonewater, B.B. (1984). Teaching problem solving: Implications from cognitive development research. AAHE Bulletin, February, 1984, 7-10.

Strange, C.C. and King, P.M. (1981). Intellectual development and its relationship to maturation during the college years. Journal of Applied Developmental Psychology, 2, 281-295.

Terman, L.M. (1973). Concept mastery test: Manual. New York: Psychological Corporation.

Welfel, E.R. (1979). The development of reflective judgment: its relationship to year in college, major field, academic performance, and satisfaction with major among college students. (Doctoral dissertation, University of Minnesota, 1979). Dissertation Abstracts International, 40, 4949A.

Welfel, E.R. (1982). How students make judgments: Do educational level and academic major make a difference? Journal of College Student Personnel, 23, 490-497.

Welfel, E.R. and Davison, M.L. (1983). Four years later: A longitudinal study of the development of reflective judgment during the college years. Technical Report, Spencer Foundation.

Wood, P.K. (1980). An analysis of the structural relationships between two tests of critical thinking and reflective judgment. Unpublished masters thesis, University of Iowa, Iowa City.

Wood, P.K. (1983). Inquiring systems and problem structure: Implications for cognitive development. Human Development, 26, 249-265.

3/A

Table 1

Reflective Judgment - Development of Epistemic Cognition

Stage	How Certain is Knowledge?	How is Knowledge Gained?	How are Beliefs Justified?
1	Absolutely certain.	By direct observation.	Beliefs are a direct reflection of reality. No need to justify them.
2	Absolutely certain but not immediately available.	By direct observation and via what authorities say is true.	Direct observation or via authorities.
3	Absolutely certain about some things; temporarily uncertain about others.	Via authorities in some areas; through our own biases when knowledge is uncertain.	Via authorities in some areas; via what feels right at the moment where knowledge is uncertain.
4	No certainty because of situational variables (e.g., data lost over time).	Via our own and others' biases, data, and logic.	Via idiosyncratic evaluations of evidence and unevaluated beliefs.
5	No certainty except via personal perspectives within a specific context.	Via evidence and rules of inquiry appropriate for the context.	By rules of inquiry for a particular context.
6	Some personal certainty about beliefs based on evaluations of evidence on different sides of the question.	Via personal assessment of arguments and data, via evaluated opinions of experts.	Via generalized rules of inquiry, personal evaluations that apply across context, evaluated views of experts.
7	Certainty that some knowledge claims are better or more complete than others although they are open to re-evaluation.	Via process of critical inquiry or synthesis.	As more or less reasonable conjectures about reality or the world based on an integration and evaluation of data, evidence and/or opinion.

BEST COPY AVAILABLE

Table 2

Mean Reflective Judgment Scores by Educational Level for Traditionally Aged Students Across Twelve Studies

	Educational Level											
	High School				College						Graduate	
Reflective Judgement	Junior		Senior		Freshmen		Sophomore		Junior		Senior		Beginning		Advanced	
Mean Score	$\bar{x}$	S[a]	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S
2.0[b]																
2.5																
3.0	2.77 (1)				3.23 (10)											
3.5			3.40 (8)		3.31 (5)		3.40 (8)		3.54 (11)		3.56 (10)					
					3.35 (11)				3.65 (1)		3.70 (8)					
					3.62 (9)				3.73 (7)							
					3.74 (6)											
4.0					3.81 (12)				3.93 (2)		3.78 (3)		4.00 (8)			
											3.97 (7)					
											4.00 (9)					
											4.08 (5)					
4.5											4.27 (6)		4.60 (4)			
5.0											4.98 (12)				4.76 (5)	
															4.96 (4)	
5.5															5.67 (1)	
6.0																
6.5																
7.0																
Overall Mean	2.77		3.40		3.60		3.40		3.68		4.08		4.24		5.04	
N	20		30		216		30		110		207		49		80	

[a]Sample source - Numbers correspond to studies, listed below

[b]A mean score of 2.0 includes the range 1.75 - 2.24; 2.5 includes 2.25 - 2.74, etc.

NOTE: All differences between groups within studies are significant (p .05) except for Brabeck's (1982) high school seniors and college sophomores.

Sample Source

(1) Kitchener & King (1981), n = 20 per group, Minnesota high school and Un. of Minnesota students.

(2) King & Parker (1978), n = 20. This sample is matched to the Kitchener & King (1981) sample of college juniors; mean score is significantly different from the scores of the high school and graduate students in the Kitchener & King study.

(3) Shoff (1979), n = 14, Un. of Utah students.

(4) Lawson (1980), n = 20 per group, Un. of Minnesota students.

(5) Mines (1980), n = 20 freshmen, 40 seniors, 40 advanced doctoral students, Un. of Iowa students.

(6) Strange & King (1981), n = 16 per group, Un. of Iowa students.

(7) Hayes (1981), n = 30 per group, Un. of Utah students.

(8) Brabeck (1982), n = 30 high school seniors, 30 college sophomores, 30 college seniors, and 29 beginning graduate students, private Catholic high school and college students.

(9) Welfel (1982), n = 32 per group, Un. of Minnesota students.

(10) Clatfelter (1982), n = 16 per group, Utah State Un. students.

(11) Schmidt (1983), n = 40 per group, ~~Utah State~~ Un. students.

(12) Griffith (1984), n = 92 freshmen, 29 seniors, Davidson College students.

BEST COPY AVAILABLE

Table 3

Longitudinal Changes in Reflective Judgment Mean Score by Educational Level over Five Studies

Study[a]	n[b]	Time 1			Time 2		Time 3	
		Educational Level	Year	RJ $\bar{x}$	Year	RJ $\bar{x}$	Year	RJ $\bar{x}$
1	17	high school juniors	1977	2.79	1979	3.61		
	27	college juniors	1977	3.75	1979	4.18		
	15	advanced graduate students	1977	6.03	1979	6.26		
2	14	high school juniors	1977	2.83	1979	3.63	1983	4.99
	22	college juniors	1977	3.72	1979	4.10	1983	4.88
	12	advanced graduate students	1977	6.15	1979	6.33	1983	6.27
3	25	high school seniors	1979	3.40	1980	3.56		
4	25	college freshmen	1978	3.62	1980	4.19		
5	34	college freshmen (traditional age)	1979	3.37	1982	3.62		
	11	college freshmen (non-traditional age)	1980	3.56	1982	3.58		

[a]1 - King, Kitchener, Davison, Parker & Wood (1983); 2 - Kitchener, King & Wood (1984); 3 - Brabeck & Wood (1983); 4 - Welfel & Davison (1983); 5 - Schmidt (1983).

[b]Includes only those subjects who participated at each testing time.

Table 4

Reflective Judgment Mean Scores by Sex and Educational Level

Educational Level	n[a]	RJ Mean Score		Difference
		Males	Females	
High School Juniors	20	2.64	2.89	+.25
College Students				
Freshmen	108	3.57	3.49	-.08
Juniors	80	3.77	3.72	-.05
Seniors	104	4.34	3.98	-.36
Graduate Students				
Beginning	20	4.94	4.25	-.69
Advanced	80	5.31	4.98	-.33

Note. Includes studies by King and Parker (1978), King, Mines and Wood (1984), Kitchener and King (1981), Lawson (1980), Schmidt (1983), Shoff (1979), Strange and King (1981) and Welfel (1982).

[a]Each sample is comprised of equal numbers of males and females.

Reflective Judgment Bibliography
(8-85)

Published Articles

Brabeck, M. (1983). Critical thinking skills and reflective judgment development: Redefining the aims of higher education. Journal of Applied Developmental Psychology, 4, 23-34.

Brabeck, M. (1984). Longitudinal studies of intellectual development during adulthood: Theoretical and research models. Journal of Research and Development in Education, 17(3), 12-27.

Brabeck, M., & Welfel, E. R. (1985). Counseling theory: Understanding the trend toward eclecticism from a developmental perspective. Journal of Counseling and Development, 63(6), 343-348.

King, P. M. (1982). Developing a rational belief system: Intellectual development in the college years. Proceedings, Symposium on Student Intellectual Development, Bowling Green State University.

King, P. M. (1985, December). Choice-making in young adulthood: A developmental double- bind. Counseling and Human Development.

King, P. M., Kitchener, K. S., Davison, M. L., Parker, C. A., & Wood, P. K. (1983). The justification of beliefs in young adults: A longitudinal study. Human Development, 26, 106-116.

King, P. M., Kitchener, K. S., & Wood, P. K. (1985). The development of intellect and character: A longitudinal-sequential study of intellectual and moral development in young adults. Moral Education Forum, 10(1), 1-13.

Kitchener, K. S. (1985). The reflective judgment model: Characteristics, evidence and measurement. In R. A. Mines & K. S. Kitchener (Eds.), Social-cognitive development in young adults. New York: Praeger.

Kitchener, K. S. (1983). Cognition, metacognition and epistemic cognition: A three-level model of cognitive processing. Human Development, 26, 222-232.

Kitchener, K. S. (1983). Educational goals and contemporary models of reflective thinking. Educational Forum, 75-95.

Kitchener, K. S., & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. Journal of Applied Developmental Psychology, 2, 89-116.

Kitchener, K. S., & Kitchener, R. F. (1981). The development of natural rationality: Can formal operations account for it? In J. A. Meacham & N. R. Santelli (Eds.), Social development in youth: Structure and content. Basel: S. Karger.

Sakalys, J. A. (1984). Effects of an undergraduate research course on cognitive development. Nursing Research, 33(5), 291-295.

Schmidt, J. A. (1985). Older and wiser? A longitudinal study of the impact of college on intellectual development. Journal of College Student Personnel, 26, 388-394.

Schmidt, J. A., & Davison, M. L. (1981). Does college matter? Reflective judgment: How students tackle the tough questions. Moral Education Forum, 6, 2-14.

Schmidt, J. A., & Davison, M. L. (1983). Helping students think. Personnel and Guidance Journal, 61(9), 563-569.

Strange, C. C., & King, P. M. (1981). Intellectual development and its relationship to maturation during the college years. Journal of Applied Developmental Psychology, 2, 281-295.

Welfel, E. R. (1982). How students make judgments: Do educational level and academic major make a difference? Journal of College Student Personnel, 23(6), 490-497.

Welfel, E. R. (1982). The development of reflective judgment: Implications for career counseling of college students. Personnel and Guidance Journal, 61(1), 17-21.

Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. Human Development, 26, 249-265.

Dissertations and Theses

Brabeck, M. (1980). The relationship between critical thinking skills and the development of reflective judgment. Unpublished doctoral dissertation, University of Minnesota.

Glatfelter, M. (1982). Identity development, intellectual development, and their relationship in reentry women students. Unpublished doctoral dissertation, University of Minnesota.

Hayes, A. B. (1981). An investigation of the effect of dilemma content on level of reasoning in the reflective judgment interview. Unpublished doctoral dissertation, University of Utah.

King, P. (1977). The development of reflective judgment and formal operational thinking in adolescents and young adults. Unpublished doctoral dissertation, University of Minnesota.

Kitchener, K. S. (1977). Intellectual development in late adolescents and young adults: Reflective judgment and verbal reasoning. Unpublished doctoral dissertation, University of Minnesota.

Lawson, J. M. (1980). The relationship between graduate education and the development of reflective judgment: A function of age or educational experience. Unpublished doctoral dissertation, University of Minnesota.

Mines, R. A. (1980). An investigation of the developmental levels of reflective judgment and associated critical thinking skills. Unpublished doctoral dissertation, University of Iowa.

Sakalys, J. A. (1984). Effects of a research methods course on nursing students' research attitudes and cognitive development. Unpublished doctoral dissertation, University of Denver.

Schmidt, J. A. (1983). The intellectual development of traditionally and nontraditionally aged college students: A cross sectional study with longitudinal follow-up. Unpublished doctoral dissertation, University of Minnesota.

Shoff, S. P. (1979). The significance of age, sex and type of education on the development of reasoning in adults. Unpublished doctoral dissertation, University of Utah.

Strange, C. C. (1978). Intellectual development, motive for education and learning styles during the college years: A comparison of adult and traditional-age college students. Unpublished doctoral dissertation, University of Iowa.

Welfel, E. R. (1979). The development of reflective judgment: its relationship to year in college, academic major, and satisfaction with major among college students. Unpublished doctoral dissertation, University of Minnesota.

Wood, P. K. (1980). An analysis of the structural relationships between two tests of critical thinking and reflective judgment. Unpublished master's thesis, University of Iowa.